

Shape and pattern containment of separable permutations

Andrew Crites *

Greta Panova[†]

Gregory S. Warrington[‡]

September 6, 2011

Abstract

Every word has a *shape* determined by its image under the Robinson-Schensted-Knuth correspondence. We show that when a word w contains a separable (i.e., 3142- and 2413-avoiding) permutation σ as a pattern, the shape of w contains the shape of σ . As an application, we exhibit lower bounds for the lengths of supersequences of sets containing separable permutations.

The Robinson-Schensted-Knuth (RSK) correspondence associates to any word w a pair of *Young tableaux*, each of equal partition shape $\lambda = (\lambda_1, \lambda_2, \dots)$. We say that w has *shape* $\text{sh}(w) = \lambda$. It is natural to expect that if σ is a subsequence of w , then $\text{sh}(\sigma) \subseteq \text{sh}(w)$. However, this is not necessarily the case: If $\sigma = 2413$ and $w = 24213$, then

$$(P(\sigma), Q(\sigma)) = \left(\begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}, \begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array} \right) \quad \text{and} \quad (P(w), Q(w)) = \left(\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 2 & & \\ \hline 4 & & \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 5 \\ \hline 3 & & \\ \hline 4 & & \\ \hline \end{array} \right). \quad (1)$$

We see that $\text{sh}(w) = (3, 1, 1) \not\supseteq (2, 2) = \text{sh}(\sigma)$. The main theorem of this paper is that the inclusion does hold when σ is a *separable* permutation. Furthermore, σ need only be contained as a pattern rather than as an actual subsequence.

Theorem 1. If a word w contains a separable permutation σ as a pattern, then $\text{sh}(w) \supseteq \text{sh}(\sigma)$.

Our discovery of Theorem 1 was motivated by an application involving lower bounds for *shortest containing supersequences*. Such supersequences arise in bioinformatics [13, 14] through the design of DNA microarrays, in planning [5] and in data compression [15]. This application to supersequences is described in Section 3. Section 1 introduces the requisite notation required for the proof of Theorem 1 appearing in Section 2. Section 2.1 discusses the relationship between Greene's Theorem, separable permutations, and the contents of this paper.

1 Background and setup

Let $[n]^*$ denote the set of finite-length words on the alphabet $[n] := \{1, 2, \dots, n\}$ and let $[n]^a$ denote the subset of length- a words. The set of permutations of length n is denoted by S_n (here a subset of $[n]^n$). Permutations will be denoted by Greek letters and written in one-line notation. For example, the permutation $\tau \in S_3$ defined by $\tau(1) = 3$, $\tau(2) = 1$ and $\tau(3) = 2$ is written 312. When referring to a subsequence of a permutation τ we make no distinction between the actual subsequence and the corresponding subset of elements; the subsequence can be reconstructed by the positions in τ . The length of a word u is denoted $|u|$.

Given a word $w \in [n]^a$ and a permutation $\pi \in S_m$, $m \leq a$, we say that w *contains the pattern* π if there exist indices $1 \leq i_1 < i_2 < \dots < i_m \leq a$ such that, for all $1 \leq j, k \leq m$, $w(i_j) < w(i_k)$ if and only if

*Dept. of Mathematics, University of Washington, Seattle, WA 98195, acrites@uw.edu

[†]Dept. of Mathematics, University of California, Los Angeles Los Angeles, CA 90095 panova@math.ucla.edu

[‡]Dept. of Mathematics and Statistics, University of Vermont, Burlington, VT 05401, gwarring@cems.uvm.edu. Supported in part by National Security Agency grant H98230-09-1-0023.

$\pi(j) < \pi(k)$ and $w(i_j) > w(i_k)$ if and only if $\pi(j) > \pi(k)$. If w does not contain the pattern π , then we say w *avoids* π .

We have defined pattern containment for words w as that is the generality in which we state Theorem 1. For pattern avoidance, however, we will only need the case in which w is itself a permutation. In particular, central to this paper will be permutations that simultaneously avoid the patterns 3142 and 2413. Being 3142, 2413-avoiding is one characterization of the class of *separable* permutations (see [4]). Throughout this paper σ will denote a separable permutation with $\text{sh}(\sigma) = \lambda = (\lambda_1, \lambda_2, \dots)$.

Given a permutation $\pi \in S_n$, let $P(\pi)$ denote its *inversion poset*. $P(\pi)$ has elements $(i, \pi(i))$ for $1 \leq i \leq n$ under the partial order \prec , in which $(a, b) \prec (c, d)$ if and only if $a < c$ and $b < d$. Note that increasing subsequences in π correspond to chains in $P(\pi)$. A longest increasing subsequence of π corresponds to a maximal chain in $P(\pi)$.

Example 2. The inversion poset of 2413 is $\begin{array}{cc} (2, 4) & (4, 3) \\ | & / \\ (1, 2) & (3, 1) \end{array}$ and that of 3142 is $\begin{array}{cc} (3, 4) & (4, 2) \\ | & / \\ (1, 3) & (2, 1) \end{array}$.

Example 2 above immediately gives the following fact.

Fact 3. A permutation π is separable if and only if its inversion poset $P(\pi)$ has no (induced) subposet

isomorphic to $\begin{array}{cc} * & * \\ | & / \\ * & * \end{array}$.

We write our partitions with parts in decreasing order and make no distinction between the positive and zero parts. Given a partition $\mu = (\mu_1, \mu_2, \dots)$ of n (denoted $\mu \vdash n$), the associated *Ferrers diagram* consists of μ_i left-justified cells in the i -th row from the top. A *semistandard Young tableau of shape μ* is a filling of the cells in this diagram with positive integers such that the rows weakly increase from left to right and the columns strictly increase from top to bottom. The set of such tableaux is denoted by $\text{SSYT}(\mu)$. A tableau $T \in \text{SSYT}(\mu)$, $\mu \vdash n$, is *standard* if each number from 1 to n appears in its filling. The set of all such tableaux is denoted by $\text{SYT}(\mu)$. Given a semistandard tableau T , the *reading word* of T , $\text{rw}(T)$, is the word obtained by reading off the rows from left to right starting with the bottom row. For $\mu \vdash n$, define the *superstandard tableau* $T \in \text{SYT}(\mu)$ by filling in the rows from top to bottom. That is, by placing $1, 2, \dots, \mu_1$ in the first row, $\mu_1 + 1, \mu_1 + 2, \dots, \mu_1 + \mu_2$ in the second row, etc.

The RSK correspondence yields a bijection between $[n]^a$ and $\cup_{\mu \vdash a} \text{SSYT}(\mu) \times \text{SYT}(\mu)$ [9]. We give a brief description of how to compute the pair $(P(w), Q(w))$ to which a word $w \in [n]^a$ corresponds. Write $w = w'x$ with $w' \in [n]^{a-1}$. By induction, we know that w' maps to some pair $(P(w'), Q(w'))$. We *row insert* x in the first row of $P(w')$ as follows: If $x = x_1$ is greater than or equal to all elements in this row, place x_1 at the end of the row. Otherwise, find the leftmost entry, x_2 , in the row that is strictly greater than x_1 . Place x_1 in this position and “bump” x_2 to be inserted into the next row. This process generates some finite sequence x_1, \dots, x_k of bumped elements and ends by adding x_k at the end of the k -th row, creating a new semistandard tableau $P(w)$. Set $Q(w)$ to have an a in the new box (end of row k) created by the bumping process. The *shape* of w , $\text{sh}(w)$, is the shape of $P(w)$ (or, equivalently, of $Q(w)$).

Example 4. The permutation $\pi = 7135264$ contains the pattern 4231 but avoids 3412. Under the RSK

correspondence, $w = 2214312$ maps to $(P(w), Q(w)) = \left(\begin{array}{|c|c|c|} \hline 1 & 1 & 2 \\ \hline 2 & 2 & 3 \\ \hline 4 & & \\ \hline \end{array}, \begin{array}{|c|c|c|} \hline 1 & 2 & 4 \\ \hline 3 & 5 & 7 \\ \hline 6 & & \\ \hline \end{array} \right)$ with $\text{rw}(P(w)) = 4223112$.

Finally, the superstandard tableau of shape $(3, 3, 2)$ is $\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline 4 & 5 & 6 \\ \hline 7 & 8 & \\ \hline \end{array}$.

2 Proof of Theorem 1

Many properties of a word w translate to natural properties of the associated tableaux. For example, the length of the longest weakly increasing subsequence of w equals λ_1 . In fact, Greene's Theorem [7] gives a much more precise correspondence.

Theorem 5 (Greene's Theorem). Let w be a word of shape λ . For any $d \geq 0$ the sum $\lambda_1 + \dots + \lambda_d$ equals the maximum number of elements in a disjoint union of d weakly increasing subsequences of w .

In order to prove Theorem 1, we will combine the insight afforded by Greene's Theorem with the ability to exchange collections of disjoint increasing subsequences with other collections for which the number of intersections has, in a certain sense, been reduced. Lemma 6, which is the only place separability explicitly appears in our proof, allows us to perform these exchanges.

Lemma 6. Let u , w , and w' be increasing subsequences of a separable permutation σ . Assume further that w and w' are disjoint. Then there exist two disjoint increasing subsequences α and β , such that $\alpha \cup \beta = w \cup w'$ and $\alpha \cap u = \emptyset$.

Proof. Write $u = u_0 \cup u_1$ with $u_0 \cap (w \cup w') = \emptyset$ and $u_1 \subset w \cup w'$. Since $\alpha \subset w \cup w'$, the requirement that α and u_1 be disjoint ensures that α and u are disjoint as well. Hence, without loss of generality, we may restrict our attention in the proof to the case in which $u \subset w \cup w'$.

Let $x = w \cup w'$ be the subsequence of σ containing both w and w' .

We give first a short proof by contradiction. Consider the inversion poset $P(\sigma)$ of the separable permutation σ . Increasing subsequences are in correspondence with chains and we will regard them as such. Assume there is no chain $\beta \subset (w \cup w')$, such that $u \subset \beta$ and $(w \cup w') \setminus \beta$ is also a chain. Let $\gamma \subset (w \cup w')$ be some maximal chain, such that $u \subset \gamma$. Then there exist two incomparable points $x, y \in (w \cup w') \setminus \gamma$, i.e. $x \not\leq y$ and thus belonging to the two different chains, e.g. $x \in w$, $y \in w'$. By maximality, $x \cup \gamma$, $y \cup \gamma$ are not chains. Hence there exist $a, b \in \gamma$ for which $x \not\leq a$, $y \not\leq b$, so we must have $a \in w'$ and $b \in w$. Assume $a \succ b$, then we must have $x \succ b$ and $y \prec a$. We have $\begin{array}{c} x \\ \diagdown \\ b \end{array} \begin{array}{c} a \\ \diagup \\ y \end{array}$ with $x \not\leq a$, $x \not\leq y$, $y \not\leq b$. This is a subposet of

$P(\sigma)$ isomorphic to $\begin{array}{c} * \\ \diagdown \\ * \end{array} \begin{array}{c} * \\ \diagup \\ * \end{array}$, contradicting Fact 3.

We now give a constructive proof, which allows us to find α and β . Let $z = w \cup w'$ be the subsequence of σ containing both w and w' . We can assume that u is a subsequence of z . First note that, since z is a shuffle of two increasing disjoint words, z also avoids the pattern 321. Also, since u is a subsequence of z , there exist indices $i_1 < i_1' < \dots < i_\ell$ such that $u_j = z_{i_j}$ for each $1 \leq j \leq \ell$. It will be convenient to augment our sequences by prepending a $u_0 = z_0 < \min\{z_i\}_{1 \leq i \leq n}$ and appending a $u_{\ell+1} = z_{n+1} > \max\{z_i\}_{1 \leq i \leq n}$.

For each $1 \leq j \leq \ell$, let β^j be the sequence of left-to-right maxima from $z_{i_j} \dots z_{i_{j+1}-1}$ whose values are greater than or equal to u_j and less than u_{j+1} . Define β^0 analogously except with values *greater than* u_0 and less than u_1 . Then, $\beta = \beta^0 \dots \beta^\ell$ is, by construction, an increasing subsequence of z . (Note that β does not include u_0 or $u_{\ell+1}$.)

We now need to show that $\alpha = z \setminus \beta$ is increasing. Suppose not. Then there exists some j such that $\alpha_j = z_a > \alpha_{j+1}$. Let m be the unique value such that $i_m < a < i_{m+1}$. We now split into cases in order to argue that z must contain one of the three patterns 321, 3142 or 2413.

1. Suppose $\alpha_j > u_{m+1}$. This implies $m < \ell$ (and hence that u_{m+1} is an element of z). We argue according to the region in which the point α_{j+1} lies (see Figure 1).
 - A) Then $\alpha_j \alpha_{j+1} u_{m+1}$ forms a 321 pattern.
 - B) Since α_{j+1} is not a left-to-right maximum, there must be some element β_k lying to the northwest of α_{j+1} yet below u_{m+1} . If β_k lies to the left of α_j , then $\beta_k \alpha_j \alpha_{j+1} u_{m+1}$ forms a 2413 pattern. Otherwise, $\alpha_j \beta_k \alpha_{j+1}$ forms a 321 pattern.

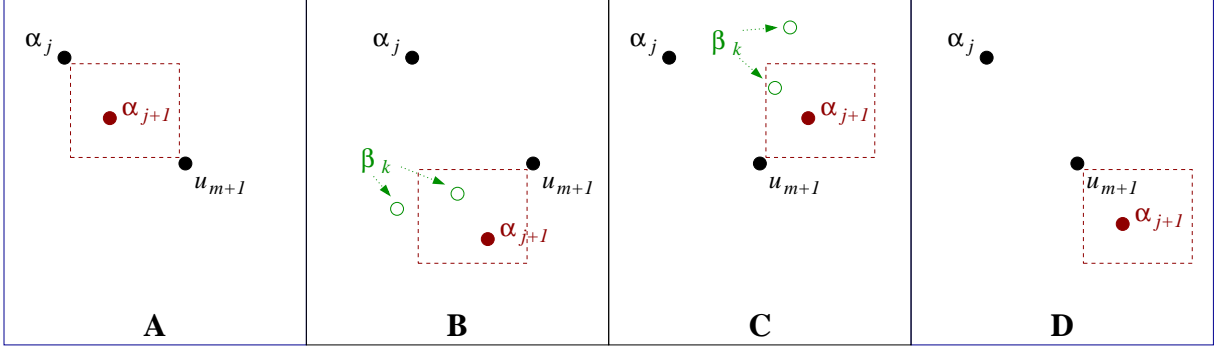


Figure 1: The cases in the proof of Lemma 6 when $\alpha_j > u_{m+1}$. Points are labeled by their y -values.

- C) As above, there must be some element β_k lying to the northwest of α_{j+1} yet to the right of u_{m+1} . If β_k lies above α_j , then $\alpha_j u_{m+1} \beta_k \alpha_{j+1}$ forms a 3142 pattern. Otherwise, $\alpha_j \beta_k \alpha_{j+1}$ forms a 321 pattern.
- D) Then $\alpha_j u_{m+1} \alpha_{j+1}$ forms a 321 pattern.
2. Suppose $\alpha_j < u_{m+1}$. Since α_j is not a left-to-right maximum, there must be some element β_k (possibly u_m) lying to the northwest of α_j . Hence $\beta_k \alpha_j \alpha_{j+1}$ forms a 321 pattern.

□

Example 7. Figure 2 illustrates the sequences α and β that arise from the construction of Lemma 6. The two original sequences shuffled together are connected by dotted lines. The elements of u are illustrated by open circles. The boxes indicate the regions in which the elements of β (other than those of u itself) are required to lie. Finally, the sequence β is connected by the thick, dashed line.

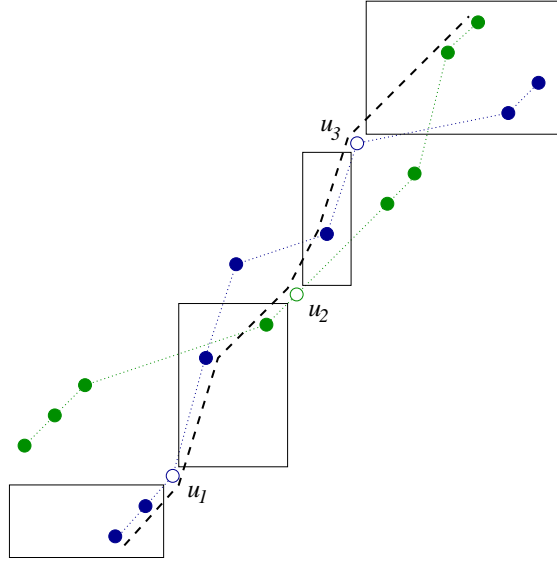


Figure 2: Example application of Lemma 6.

Proposition 8. Let $k \geq 0$ and s^1, \dots, s^k be disjoint (possibly empty) increasing subsequences of the separable permutation σ . Then there exists an increasing subsequence u^{k+1} , disjoint from each s^i , such that $|u^{k+1}| \geq \lambda_{k+1}$.

Proof. Let $V = \{v^1, \dots, v^{k+1}\}$ be any collection of $k+1$ disjoint, increasing subsequences of σ of maximum total length. Let $m = m(V)$ be the maximum index such that v^j is disjoint from s^i for all $1 \leq i < m \leq j$. We prove by induction that m can be taken to be $k+1$.

So, assume $1 \leq m < k+1$. At least one v^j with $j > m$ must intersect s^m , otherwise m would be larger than it is. But repeated application of Lemma 6 to the elements of $\{v^m, \dots, v^{k+1}\}$ produces a new set $V' = \{v^1, \dots, v^{m-1}, \tilde{v}^m, \dots, \tilde{v}^{k+1}\}$ in which $j = m$ is the only value for which \tilde{v}^j intersects s^m . Hence $m(V') \geq m(V) + 1$, completing the induction step.

By the preceding paragraph, we may assume that v^{k+1} is disjoint from all of the s^i . Since the elements of V are of maximum total length, $|v^1| + \dots + |v^{k+1}| = \lambda_1 + \dots + \lambda_{k+1}$. If $|v^{k+1}|$ were less than λ_{k+1} , then v^1, \dots, v^k would have total length greater than $\lambda_1 + \dots + \lambda_k$. This is impossible. Hence $|v^{k+1}| \geq \lambda_{k+1}$ as desired. \square

Example 9. Consider $x = 10652438ba97$ (where we use a for 10 and b for 11). The shape of x is $(5, 3, 2, 2)$. Suppose we have $u^1 = 0248b$, $u^2 = 167$ and $u^3 = 5a$ and wish to find a disjoint increasing subsequence u^4 of length 2. We could, of course, simply use the remaining two elements, 3 and 9. However, in order to illustrate the proofs of Proposition 8 and Theorem 11, we show how to generate this sequence from an arbitrarily chosen 4-tuple of disjoint increasing subsequences of maximum total length: $V = \{68b, 049, 237, 15a\}$.

So, set $k = 3$ and $s^j = u^j$ for $j \in \{1, 2, 3\}$. Consider the argument of Proposition 8. $m = m(V) = 1$. Let $u = u^1$, $w = 68b$ and $w' = 049$. Lemma 6 yields $\alpha = 69$ and $\beta = 048b$. Applying the lemma again with $w = 048b$ and $w' = 237$ yields $\alpha = 37$ and $\beta = 0248b$. This produces the new 4-tuple $V' = \{0248b, 69, 37, 15a\}$ with $m(V') = 2$.

Now set $u = u^2$. Once again, an application of the lemma with $w = 69$ and $w' = 15a$ yields $\alpha = 59$ and $\beta = 16a$, while a following application to $w = 16a$ and $w' = 37$ yields $\alpha = 3a$ and $\beta = 167$. This produces the new 4-tuple $V'' = \{0248b, 167, 59, 3a\}$ with $m(V'') = 3$.

A final application of the lemma with $u = u^3 = 5a$, $w = 59$ and $w' = 3a$ yields the sought for $u^4 = 39$.

Proof of Theorem 1. Let $\text{sh}(w) = \mu = (\mu_1, \mu_2, \dots)$. Let σ' be any subsequence of w in the same relative order as the elements of σ ; i.e., w contains σ at the positions of σ' . By Greene's Theorem applied to w , for any $k \geq 1$ there exist k disjoint increasing subsequences w^1, \dots, w^k with $|w^1| + \dots + |w^k| = \mu_1 + \dots + \mu_k$. The intersection $\sigma' \cap w^i$ induces a subsequence of σ we denote by s^i . These s^i are then k disjoint increasing subsequences of σ . By Proposition 8, there is an increasing subsequence u of σ , disjoint from the s^i s, with length at least λ_{k+1} . The mapping $\sigma \mapsto \sigma'$ induces a corresponding map of u to a subsequence u' of w . It follows then that u' is disjoint from each w^i as well. Then w^1, \dots, w^k, u' are $k+1$ disjoint increasing subsequences in w . By Greene's Theorem,

$$|w^1| + \dots + |w^k| + |u'| \leq \mu_1 + \dots + \mu_k + \mu_{k+1}.$$

Hence $|u'| \leq \mu_{k+1}$. We also know by construction that $\lambda_{k+1} \leq |u| = |u'|$. Combining these equalities and running over all k yields $\lambda \subseteq \mu$ as desired. \square

2.1 Relationship to Greene's Theorem

Greene's Theorem only tells us about the maximum *sum* of lengths of disjoint increasing sequences. It is *not* generally true that one can find d disjoint increasing subsequences u^1, u^2, \dots, u^d of w with u^i of length λ_i for each i . In other words, the shape of a word does not tell you the lengths of the subsequences in a set of d disjoint increasing subsequences of maximum total length; it just tells you the maximum total length.

Example 10. Consider the permutation $w = 236145$ of shape $(4, 2)$. The only increasing subsequence of length four is 2345. However, the remaining two entries appear in decreasing order. Greene's Theorem tells us that we should be able to find two disjoint increasing subsequences of total length 6. Indeed, 236 and 145 work.

Nonetheless, such a collection of subsequences $\{u^i\}$ does exist when σ is a separable permutation.

Proposition 11. Let σ be a separable permutation of shape λ . For any $d \geq 1$, there exist d disjoint, increasing subsequences u^1, \dots, u^d such that the length of each u^i is given by λ_i .

Theorem 1 and Proposition 11 are superficially similar. We have already shown how Theorem 1 follows from Proposition 8 (and Greene's Theorem). Proposition 11 follows even more immediately.

Proof of Proposition 11. We can construct such a sequence via d applications of Proposition 8. In particular, given the u^1, \dots, u^i for some $0 \leq i < d$, produce u^{i+1} by applying the proposition with $k = i$ and $s^j = u^j$ for $1 \leq j \leq k$. \square

As pointed out to us by a referee to an earlier version of this paper, Proposition 11 has a very simple proof relying on the recursive definition of a separable permutation as one that can be built up by direct and skew sums (see [4]). However, we have been unable to follow a correspondingly direct proof of Theorem 1.

3 Supersequences

Let $B \subseteq S_n$ be a set of permutations. A word w is a *supersequence* of B if, for all $\sigma \in B$, σ is a subsequence of w . Note that for w to be a supersequence of $\{\sigma\}$, the actual entries of σ must occur (in the same order) in w ; this is in contrast to pattern containment in which we need only find elements of w in the same relative order.

Example 12. The word $w = 2214312$ is a supersequence of 132 but not of 321. In fact, w is a supersequence of the set $B = \{132, 312, 213\}$.

Let $\text{scs}_n(B)$ denote the minimum length of a supersequence of the set B . An upper bound of $\text{scs}_n(S_n) \leq n^2 - 2n + 4$ has been proven by a number of different researchers in various contexts and generalities. See in particular [1, 6, 10, 11, 12, 16]. Recently, an upper bound of $n^2 - 2n + 3$ was proven constructively for $n \geq 10$ by Zălinescu [17]. Kleitman and Kwiakowski [8] have shown that $\text{scs}_n(S_n) \geq n^2 - Cn^{7/4+\varepsilon}$ where $\varepsilon > 0$ and C depends on ε .

For certain sets B , we can construct a lower bound for $\text{scs}_n(B)$ by considering the union of $\text{sh}(\sigma)$ as σ runs over the elements of B .

Lemma 13. If T is the superstandard tableau of shape λ , then $\text{rw}(T)$ is a 2413,3142-avoiding permutation (i.e., is separable).

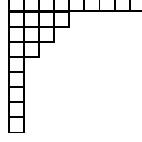
Proof. In fact, $\text{rw}(T)$ avoids the pattern 213: For $i < j$, the entries in row j are greater than, and precede, the entries in row i . Hence $\text{rw}(T)$ avoids both 2413 and 3142. \square

Fix $k > 0$ and $B = \{\sigma_1, \dots, \sigma_k\}$ with each σ_i separable. It follows then from Theorem 11 that for any supersequence w of B , $\text{sh}(w) \supseteq \cup_i \text{sh}(\sigma_i)$. Hence, if we choose the σ_i so that the Ferrers diagrams of shapes $\text{sh}(\sigma_i)$ overlap as little as possible, we force any supersequence w to be relatively long.

Example 14. Let $n = 9$ and $k = 5$. Choose the permutations $B = \{\sigma_1, \dots, \sigma_5\}$ as

$$\begin{aligned} \sigma_1 &= 123456789, & \text{sh}(\sigma_1) &= (9), \\ \sigma_2 &= 678912345, & \text{sh}(\sigma_2) &= (5, 4), \\ \sigma_3 &= 789456123, & \text{sh}(\sigma_3) &= (3, 3, 3), \\ \sigma_4 &= 978563412, & \text{sh}(\sigma_4) &= (2, 2, 2, 2, 1), \\ \sigma_5 &= 987654321, & \text{sh}(\sigma_5) &= (1, 1, 1, 1, 1, 1, 1, 1, 1). \end{aligned}$$

The union of the corresponding Ferrers diagrams is



; we see that $|\cup_{i=1}^5 \text{sh}(\sigma_i)| = 23$. A computer

search provides the length-23 supersequence

$$69787596543123456789123,$$

thereby showing that this bound is optimal.

Let $\mu(n)$ be the Ferrers diagram obtained by taking the union of all Ferrers diagrams of size n .

Proposition 15. Let $\tau(i)$ denote the number of divisors of i . Then $|\mu(n)| = \sum_{i=1}^n \tau(i)$ and the number of corners (i.e., distinct row lengths of $\mu(n)$) is given by $\lfloor \sqrt{4n+1} \rfloor - 1$.

Proof. For each divisor d of n , the shape $\overbrace{(n/d, \dots, n/d)}^{d \text{ times}}$ will be contained in $\mu(n)$. Furthermore, each cell $(d, n/d)$ will be a corner that is not part of $\mu(n-1)$. The result $|\mu(n)| = \sum_{i=1}^n \tau(i)$ then follows by induction. (In fact, the nested sequence of Ferrers diagrams $\mu(1) \subset \mu(2) \subset \dots \subset \mu(n)$ can be thought of as a semistandard Young tableau of shape $\mu(n)$ in which the label i occurs $\tau(i)$ times.)

We now prove that the number of corners of $|\mu(n)|$ is given by $\lfloor \sqrt{4n+1} \rfloor - 1$. Let k be the largest integer for which a $k \times k$ square is contained in the diagram of $\mu(n)$, that is, k is the number of cells on the main diagonal in $\mu(n)$. We have that $k^2 \leq n$. The cell (k, k) is a corner of $\mu(n)$ if and only if $k(k+1) > n$, i.e. $\underbrace{(k, \dots, k)}_{k+1}$ is not contained in any diagram of size n . We claim that the rows $1, \dots, k$ of $\mu(n)$ will each contain

a corner of $\mu(n)$. This is trivially true for the first row. For $1 < i \leq k$, row i ends at $(i, \lfloor n/i \rfloor)$ while the row above ends at $(i-1, \lfloor n/(i-1) \rfloor)$. Since $k \leq \sqrt{n}$, $\lfloor n/i \rfloor < \lfloor n/(i-1) \rfloor$ and we have a corner in the i -th row as claimed. The same argument holds for the first k columns, so the total number of corners is $2k-1$ if (k, k) is a corner and $2k$ otherwise. We have that $(2k+1)^2 \geq 4n+1 \geq 4k^2+1$, with the first inequality strict if and only if (k, k) is a corner. Hence the number of corners is indeed $\lfloor \sqrt{4n+1} \rfloor - 1$. \square

It is a standard fact that $\sum_{i=1}^n \tau(i) \sim n(\ln n + 2\gamma + \dots)$ (see, e.g., [3, Theorem 3.3]). Hence, for any n we can find $\lfloor \sqrt{4n+1} \rfloor - 1$ permutations whose supersequence is of length at least $n(\ln n + 2\gamma + \dots)$. Compare this with $n!$ permutations having a supersequence of length $O(n^2)$.

4 Acknowledgments

The authors would like to thank Dan Archdeacon, Jeff Buzas and Jeff Dinitz for useful conversations leading to this paper.

References

- [1] Leonard Adleman, *Short permutation strings*, Discrete Math. **10** (1974), 197–200.
- [2] Michael Albert, *Young classes of permutations*, arXiv:1008.4615v2 [math.CO].
- [3] Tom M. Apostol, *Introduction to analytic number theory*, Springer-Verlag, New York, 1976, Undergraduate Texts in Mathematics.
- [4] Prosenjit Bose, Jonathan F. Buss, and Anna Lubiw, *Pattern matching for permutations*, Inform. Process. Lett. **65** (1998), no. 5, 277–283.
- [5] D. Foulser, M. Li and Q. Yang, *Theory and algorithms for plan merging*, Artificial Intelligence, **57** (1992) no. 2-3, 143–181.

- [6] G. Galbiati and F. P. Preparata, *On permutation-embedding sequences*, SIAM J. Appl. Math. **30** (1976), no. 3, 421–423.
- [7] Curtis Greene, *An extension of Schensted’s theorem*, Advances in Math. **14** (1974), 254–265.
- [8] D. J. Kleitman and D. J. Kwiatkowski, *A lower bound on the length of a sequence containing all permutations as subsequences*, J. Combinatorial Theory Ser. A **21** (1976), no. 2, 129–136.
- [9] Donald E. Knuth, *Permutations, matrices, and generalized Young tableaux*, Pacific J. Math. **34** (1970), 709–727.
- [10] P. J. Koutas and T. C. Hu, *Shortest string containing all permutations*, Discrete Math. **11** (1975), 125–132.
- [11] S. P. Mohanty, *Shortest string containing all permutations*, Discrete Math. **31** (1980), no. 1, 91–95.
- [12] M. C. Newey, *Notes on a problem involving permutations as subsequences*, Technical Report **CS-TR-73-340** (1973).
- [13] K. Ning and H. W. Leong, *Towards a better solution to the shortest common supersequence problem: the deposition and reduction algorithm*, BMC Bioinformatics **7** (2006), Suppl 4:S12.
- [14] S. Rahmann, *The shortest common supersequence problem in a microarray production setting*, Bioinformatics (2003), Suppl 2:ii, 156–61.
- [15] M. Rodeh, V. R. Pratt and S. Even, *Linear algorithm for data compression via string matching*, J. ACM **28** (1981), no. 1, 16–24.
- [16] Carla Savage, *Short strings containing all k -element permutations*, Discrete Math. **42** (1982), no. 2-3, 281–285.
- [17] Eugen Zălinescu, *Shorter strings containing all k -element permutations*, Inform. Process. Lett. **111** (2011), no. 12, 605–608.